Nonlocal Filters for Removing Multiplicative Noise

Tanja Teuber[1] and Annika Lang[2]

[1] Department of Mathematics, University of Kaiserslautern, Germany
tteuber@mathematik.uni-kl.de
[2] Seminar for Applied Mathematics, ETH Zurich, Switzerland
annika.lang@sam.math.ethz.ch

Abstract. In this paper, we propose nonlocal filters for removing multiplicative noise in images. The considered filters are deduced in a weighted maximum likelihood estimation framework and the occurring weights are defined by a new similarity measure for comparing data corrupted by multiplicative noise. For the deduction of this measure we analyze a probabilistic measure recently proposed for general noise models by Deledalle et al. and study its properties in the presence of additive and multiplicative noise. Since it turns out to have unfavorable properties facing multiplicative noise we propose a new similarity measure consisting of a density specially chosen for this type of noise. The properties of our new measure are examined theoretically as well as by numerical experiments. Afterwards, it is applied to define the weights of our nonlocal filters and different adaptations are proposed to further improve the results. Throughout the paper, our findings are exemplified for multiplicative Gamma noise. Finally, restoration results are presented to demonstrate the good properties of our new filters.

1 Introduction

In 2005, Buades et al. introduced the well-known *nonlocal (NL) means filter* [3]. For the restoration this filter uses information gained by comparing various image regions, so-called patches, with each other. In detail, for a discrete image $f \in \mathbb{R}^{m,n}$, $N = mn$ with pixels f_i, $i = 1, \ldots, N$, the restored pixels are set to be

$$\widetilde{u}_i = \frac{1}{C_i} \sum_{j=1}^{N} w_{NL}(i,j) f_j \qquad \text{with} \;\; C_i := \sum_{j=1}^{N} w_{NL}(i,j). \tag{1}$$

If the image patches with centers f_i, f_j are given by f_{i+I}, resp. f_{j+I} for I denoting an appropriate index set, then the weights are given by

$$w_{NL}(i,j) \;=\; \exp\left(-\tfrac{1}{h} \sum_{k \in I} g_k |f_{i+k} - f_{j+k}|^2 \right).$$

Here, $h > 0$ controls the amount of filtering. The vector $g = (g_k)_{k \in I}$ represents usually a sampled two dimensional Gaussian kernel with mean zero and standard deviation a, which steers the influence of neighboring pixels on the weight.

A.M. Bruckstein et al. (Eds.): SSVM 2011, LNCS 6667, pp. 50–61, 2012.

This filter has been extensively studied in the past five years and further improved in various directions. An overview is for example given in [4]. One improvement was that several authors proposed different approaches to adapt the NL means filter to noise statistics. Kervrann et al. proposed the so-called *Bayesian NL means filter* [10], which was applied for the removal of speckle noise in ultrasound images in [5]. For Rician noise an approach was presented in [16]. Another relative of the original NL means filter in a probabilistic framework was proposed by Deledalle et al. in [6]. Their approach involved a new noise dependent similarity measure for the patch comparison and was demonstrated to perform well for images corrupted by additive Gaussian noise, noise following a Nakagami-Rayleigh distribution as well as Poisson noise studied in [7].

The aim of this paper is to present nonlocal filters for removing multiplicative noise. To exemplify our results we concentrate on multiplicative Gamma noise. Note that all missing proofs and further examples including different types of noise can be found in [15]. In Section 2 we start by defining our filters by maximum likelihood estimation. For the weight definition we propose a new similarity measure specially designed for comparing data corrupted by multiplicative noise. To obtain this measure we analyze the similarity measure of [6] in the framework of conditional densities in Section 3 and study its properties facing additive and multiplicative noise. Since it turns out to be well suited for additive noise, but to have unfavorable properties for multiplicative noise, we deduce our new measure by logarithmically transformed random variables in Section 4. The advantages of our measure are shown theoretically and by numerical experiments. In Section 5, we consider variants of the weight definition, which further improve the results. Finally, the very good performance of our novel nonlocal filters is demonstrated for images corrupted by multiplicative Gamma noise in Section 6.

2 Nonlocal Filters for Multiplicative Noise

As proposed in [6,12], we will deduce our nonlocal filters by weighted maximum likelihood estimation. Throughout this paper, all random variables are supposed to be continuous and defined on a fixed probability space $(\Omega, \mathcal{F}, P)$. Moreover, for a random variable X and a constant $c \in \mathbb{R}$ we denote by p_{cX} the density of the random variable cX. For $x \in \mathbb{R}$ with $p_X(x) > 0$, the *conditional density* of Y given $X = x$ is defined by $p_{Y|X}(\cdot \,|\, x) := \frac{p_{Y,X}(\cdot, x)}{p_X(x)}$, see, e.g., [9, p. 104]. Now, assume that for $i = 1, \ldots, N$ all noisy image pixels f_i are realizations of independent random variables F_i and the corresponding initial noise free pixels u_i are realizations of independent and identically distributed (i.i.d.) random variables U_i. Moreover, suppose that all f_i are corrupted by the same noise model with equal parameters. Then, we define our restored pixels by

$$\widetilde{u}_i := \operatorname*{argmax}_t \sum_{j=1}^{N} w(i,j) \, \ln p_{F_j|U_j}(f_j \,|\, t) \quad \text{s.t. } p_{U_1}(t) = \cdots = p_{U_N}(t) > 0, \quad (2)$$

where $w(i,j) \in [0,1]$ is ideally one if $u_i = u_j$ and zero otherwise. If $w = w_{NL}$, we obtain for additive Gaussian noise and positive p_{U_i} that $\widetilde{u}_i$ is given by (1) as outlined in [6]. For the case of *multiplicative Gamma noise*, we assume that

$$F_i = U_i V_i, \qquad \text{with} \;\; p_{U_i}(t) = 0 \quad \forall\, t < 0,\; i = 1, \ldots, N, \tag{3}$$

where all V_i are continuous random variables with density

$$p_{V_i}(v) = \frac{L^L}{\Gamma(L)} v^{L-1} \exp(-Lv)\, 1_{\mathbb{R}_{\geq 0}}(v), \quad L \geq 1 \tag{4}$$

and Γ denotes the Gamma function. Besides, all U_i, V_i are considered pairwise independent. Then, for $j = 1, \ldots, N$ and any $f_j, t > 0$ with $p_{U_j}(t) > 0$ we have

$$p_{F_j|U_j}(f_j \,|\, t) \;=\; \frac{1}{|t|}\, p_{V_j}\left(\frac{f_j}{t}\right) \;=\; \frac{L^L}{\Gamma(L)} \frac{f_j^{L-1}}{t^L} \exp\left(-L\, \frac{f_j}{t}\right). \tag{5}$$

For $f_j > 0,\; j = 1, \ldots, N$, this implies

$$\tilde{u}_i \;=\; \underset{\substack{t>0 \\ p_{U_i}(t)>0}}{\operatorname{argmax}} \sum_{j=1}^{N} w(i,j)\, \ln p_{F_j|U_j}(f_j \,|\, t) \;=\; \underset{\substack{t>0 \\ p_{U_i}(t)>0}}{\operatorname{argmin}} \sum_{j=1}^{N} w(i,j)\, (\ln(t) + \frac{f_j}{t}).$$

Similarly, $H(f,u) := \sum_{i=1}^{N} \ln(u_i) + \frac{f_i}{u_i}$ has been deduced as a data fidelity term for a variational approach to remove multiplicative Gamma noise in [1]. If $p_{U_i}(t) > 0$ for $t > 0$ or p_{U_i} is simply unknown, we omit the restriction $p_{U_i}(t) > 0$ and obtain for $f_j > 0$, $j = 1, \ldots, N$, by the first order optimality condition that

$$\tilde{u}_i = \frac{1}{C_i} \sum_{j=1}^{N} w(i,j) f_j \qquad \text{with} \;\; C_i := \sum_{j=1}^{N} w(i,j). \tag{6}$$

Hence, we get for multiplicative Gamma noise an ordinary weighted average filter like the original NL means filter in (1). Next, we would like to define the weights similarly to w_{NL}, but incorporate the statistics of the noise. By

$$w_{NL}(i,j) \;=\; \prod_{k \in I} s_{NL}(f_{i+k}, f_{j+k})^{\frac{g_k}{h}} \quad \text{with} \;\; s_{NL}(x,y) \;:=\; \exp(-|x-y|^2) \tag{7}$$

we see that $w_{NL}(i,j)$ can be written as the product of all $s_{NL}(f_{i+k}, f_{j+k})^{\frac{g_k}{h}}$, where f_{i+k}, f_{j+k} are pairs of pixels of two fix image patches. The function $s_{NL} : \mathbb{R} \times \mathbb{R} \to (0,1]$ acts as a similarity measure, where $s_{NL}(f_{i+k}, f_{j+k})$ should be close to 1 if $u_{i+k} = u_{j+k}$ and close to 0 if not. Facing additive Gaussian noise, s_{NL} is known to perform well, but it can be far from optimal for other types of noise. Hence, the challenge is now to find a suitable similarity measure for our noise model.

3 The Similarity Measure of Deledalle et al.

To measure whether $u_1 = u_2$ by noisy observations f_1, f_2, Deledalle, Denis and Tupin suggest in [6] to use a so-called *'similarity probability'* denoted by

$p(\theta_1 = \theta_2|f_1, f_2)$. In their paper, θ_i is a parameter depending deterministically on u_i and we set $\theta_i = u_i$ for $i = 1, 2$. Since in general it is not clear what the probability or even conditional density of $U_1 = U_2$ given $F_1 = f_1$, $F_2 = f_2$ is, see e.g. [9, p. 111], we start by rewriting the 'similarity probability' as a conditional density: By definition we have for $p_{F_i}(f_i) > 0$, $i = 1, 2$, that

$$p(u_1 = u_2|\, f_1, f_2) \;:=\; \int_S p_{U_1|F_1}(u\,|\,f_1)\, p_{U_2|F_2}(u\,|\,f_2)\; du \tag{8}$$

and set $S := \mathrm{supp}(p_{U_i})$. Applying the definition of the conditional density and Jacobi's Transformation Formula, see e.g., [13, p. 135f], we obtain that

$$p(u_1 = u_2|f_1, f_2) \;=\; p_{U_1-U_2|(F_1,F_2)}(0\,|\,f_1, f_2). \tag{9}$$

Besides, we have

$$p_{U_1-U_2|(F_1,F_2)}(0\,|\,f_1, f_2) = \frac{\int_S p_{U_1}(u)\, p_{U_2}(u)\, p_{F_1|U_1}(f_1\,|\,u)\, p_{F_2|U_2}(f_2\,|\,u)\; du}{p_{F_1}(f_1)\, p_{F_2}(f_2)}. \tag{10}$$

Since normally p_{U_i} is unknown, Deledalle et al. propose to neglect this density and p_{F_i}, $i = 1, 2$, on the right hand side and to consider only

$$s_{DDT}(f_1, f_2) \;:=\; \int_S p_{F_1|U_1}(f_1\,|\,u)\; p_{F_2|U_2}(f_2\,|\,u)\; du. \tag{11}$$

This measure is very close to the one investigated for block matching in [11]. To study its properties we start by considering data corrupted by additive noise.

3.1 Properties in the Presence of Additive Noise

For $i = 1, 2$ let the random variables V_i be i.d.d. and follow some noise distribution. Moreover, let f_i be corrupted by *additive noise*, i.e. $f_i := u_i + v_i$ and

$$F_i := U_i + V_i, \qquad i = 1, 2.$$

Here, v_i is a realization of V_i and all U_i, V_i, $i = 1, 2$, are considered to be pairwise independent. In this case, we can show that s_{DDT} has the following properties:

Proposition 1. *For our additive noise model with* $S = \mathrm{supp}(p_{U_i}) = \mathbb{R}$ *we have*

$$s_{DDT}(f_1, f_2) \;=\; p_{V_1-V_2}(f_1-f_2) \;=\; p_{F_1-F_2|U_1-U_2}(f_1-f_2|\,0\,), \quad f_1, f_2 \in \mathbb{R}. \tag{12}$$

Moreover, s_{DDT} *is symmetric and has the following properties:*

i) $s_{DDT}(f, f) = const$ *for all* $f \in \mathbb{R}$,
ii) $0 \le s_{DDT}(f_1, f_2) \le s_{DDT}(f, f) = p_{V_1-V_2}(0)$ *for all* $f_1, f_2, f \in \mathbb{R}$.

For the proof of this and the following propositions see [15]. The last property implies that $s_{DDT}(f_1, f_2)$ is maximal whenever $f_1 = f_2$ and that it is bounded so that it can be scaled to the interval $[0, 1]$, i.e. the range of s_{NL}. For the special case that V_i, $i = 1, 2$, are normally distributed with standard deviation σ, it follows that

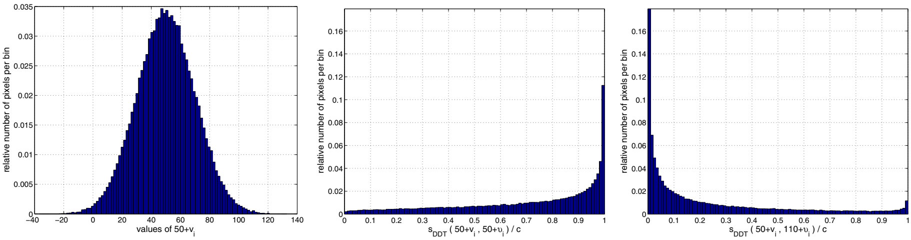

Fig. 1. *Left*: Histogram of a constant image of gray value 50 corrupted by additive Gaussian noise with $\sigma = 20$. *Middle*: Histogram of $(s_{DDT}(f_i, \widetilde{f}_i)/c)_{i=1}^N$, where f, $\widetilde{f}$ are images with gray value distributions as on the left. *Right*: Same as in the middle, but now $\widetilde{f}$ represents a constant image of gray value 110 corrupted by noise.

$$s_{DDT}(f_1, f_2) \; = \; \frac{1}{2\sqrt{\pi}\sigma} \exp\left(-\frac{|f_1 - f_2|^2}{4\sigma^2}\right) \; = \; c\,(\, s_{NL}(f_1, f_2)\,)^{\frac{1}{4\sigma^2}}$$

with $c := \max\limits_{x,y \in \mathbb{R}} s_{DDT}(x, y) = \frac{1}{2\sqrt{\pi}\sigma}$. The behavior of s_{DDT} for additive Gaussian noise is illustrated in Fig. 1. In the middle, the distribution of the values $s_{DDT}(f_i, \widetilde{f}_i)/c$ is depicted if both images f, $\widetilde{f}$ are corrupted versions of the same constant image. As expected, most values are close to 1, i.e. s_{DDT}/c detected that the corresponding noisy pixels belong to the same noise free pixel. Only a few values are close to zero, where the measure did not recognize that also these noisy pixels have the same initial gray value. On the right, where the initial gray values have been different, most values $s_{DDT}(f_i, \widetilde{f}_i)/c$ are close to zero and only few pixels are falsely detected to correspond to the same noise free pixel.

3.2 Properties in the Presence of Multiplicative Noise

Next, we want to investigate the case of multiplicative noise. We suppose that the random variables V_i, $i = 1, 2$, are i.i.d., pairwise independent with both U_i and $p_{V_i}(x) = 0$ for $x < 0$. Besides, we assume that F_i follows the multiplicative noise model (3) so that $F_i > 0$ almost surely for $i = 1, 2$. For this setting, we obtain the following properties of s_{DDT}:

Proposition 2. *For our multiplicative noise model with* $S = \mathrm{supp}(p_{U_i}) = \mathbb{R}_{\geq 0}$ *and* $f_1, f_2 > 0$ *it holds that*

$$s_{DDT}(f_1, f_2) \; = \; \int_0^\infty \frac{1}{u^2}\, p_{V_1}\left(\frac{f_1}{u}\right) p_{V_2}\left(\frac{f_2}{u}\right)\, du \; = \; p_{f_2 V_1 - f_1 V_2}(0). \qquad (13)$$

Besides, s_{DDT} *is symmetric and has the following properties:*

i) $s_{DDT}(f, f) = \frac{1}{f}\, p_{V_1 - V_2}(0)$ *for all* $f = f_1 = f_2 > 0$*,*
ii) s_{DDT} *is not bounded from above.*

These properties stand in sharp contrast to the additive case. The first property implies that s_{DDT} always considers small values $f = f_1 = f_2$ more likely to

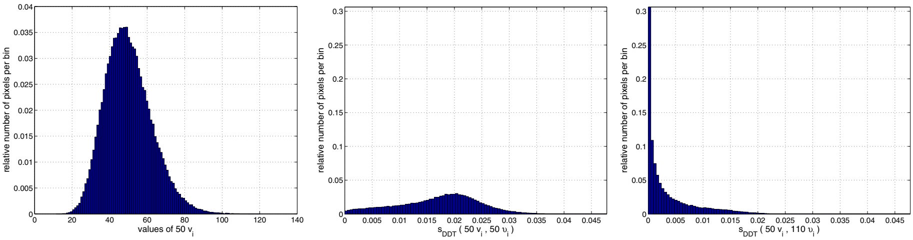

Fig. 2. *Left*: Histogram of a constant image with gray value 50 corrupted by multiplicative Gamma noise with $L = 16$. *Middle*: Histogram of $(s_{DDT}(f_i, \widetilde{f}_i))_{i=1}^N$, where f, $\widetilde{f}$ have gray value distributions as on the left. *Right*: Same as in the middle, but now $\widetilde{f}$ represents a constant image of gray value 110 corrupted by noise.

have the same initial gray value than bigger ones. Besides, the unboundedness is problematic with regard to the weight definition of our nonlocal filters, since a single pixel could get an arbitrarily large weight and dominate all others.

For multiplicative Gamma noise we obtain for $f_1, f_2 > 0$ and $S = \mathbb{R}_{\geq 0}$ that

$$s_{DDT}(f_1, f_2) \;=\; L\frac{\Gamma(2L-1)}{\Gamma(L)^2}\,\frac{(f_1 f_2)^{L-1}}{(f_1+f_2)^{2L-1}} \;\propto\; \frac{1}{f_1+f_2}\,\left(2+\frac{f_1}{f_2}+\frac{f_2}{f_1}\right)^{1-L}.$$

One may expect that for fixed f_1, s_{DDT} is maximal if $f_2 = f_1$. However, for $L > 1$ and a given value f_1 it is maximal for $f_2 = \frac{L-1}{L} f_1$. This is again in sharp contrast to the properties of s_{DDT} in the additive case. For $L = 1$ we have $s_{DDT}(f_1, f_2) = \frac{1}{f_1+f_2}$. Thus, $s_{DDT}(f_1, f_2)$ is large whenever f_1, f_2 are small. Further properties of this measure are illustrated for $L = 16$ in Fig. 2. In contrast to Fig. 1 (middle), the peak of the histogram at Fig. 2 (middle) is no longer at the largest obtained value of the measure, but at some intermediate value. This is not desirable with respect to the weight definition of a nonlocal filter, since for a large number of pixels it would not definitely determine whether the true pixels have been the same or not. Hence, s_{DDT} does not seem to be optimal for multiplicative noise.

4 A New Similarity Measure for Multiplicative Noise

To deduce a different measure for our multiplicative noise model, we consider the transformed random variables $\widetilde{F}_i = \ln(F_i)$, $\widetilde{U}_i = \ln(U_i)$, $\widetilde{V}_i = \ln(V_i)$, where

$$\widetilde{F}_i = \ln(F_i) \;=\; \ln(U_i V_i) \;=\; \widetilde{U}_i + \widetilde{V}_i, \qquad i = 1, 2.$$

The new random variables $\widetilde{F}_i$ follow an additive noise model now and the supports of $p_{\widetilde{U}_i}$, $p_{\widetilde{V}_i}$ may be the whole of $\mathbb{R}$. By computing (9) for these new random variables we can show the following:

Lemma 1. *For $f_1, f_2 > 0$ with $p_{F_i}(f_i) > 0$ and $\widetilde{S} = \operatorname{supp}(p_{\widetilde{U}_i})$ it holds that*

$$p_{\widetilde{U}_1-\widetilde{U}_2|(\widetilde{F}_1,\widetilde{F}_2)}(0 \mid \ln(f_1), \ln(f_2)) = p_{\frac{U_1}{U_2}|(F_1,F_2)}(1 \mid f_1, f_2). \tag{14}$$

Compared to (9), we have replaced $U_1 - U_2 = 0$ by $U_1/U_2 = 1$ now. Next, we use (10) for the transformed variables and omit $p_{\widetilde{U}_i}$, $p_{\widetilde{F}_i}$, $i = 1, 2$. Supposing that $\widetilde{S} = \mathbb{R}$, i.e. $S = \mathbb{R}_{\geq 0}$, and using (12) for the right hand side, we thus obtain

$$\int_{\widetilde{S}} p_{\widetilde{F}_1|\widetilde{U}_1}(\ln(f_1)\,|\,t)\; p_{\widetilde{F}_2|\widetilde{U}_2}(\ln(f_2)\,|\,t)\; dt \;=\; p_{\widetilde{V}_1 - \widetilde{V}_2}(\ln(f_1) - \ln(f_2)).$$

Defining our new similarity measure by

$$s(f_1, f_2) := p_{\widetilde{V}_1 - \widetilde{V}_2}(\ln(f_1) - \ln(f_2)) \;=\; p_{\widetilde{F}_1 - \widetilde{F}_2|\widetilde{U}_1 - \widetilde{U}_2}(\ln(f_1) - \ln(f_2)\,|\,0), \quad (15)$$

it has the following properties similar to s_{DDT} for $S = \mathbb{R}$ in the additive case:

Proposition 3. *For our multiplicative noise model and $f_1, f_2 > 0$ it holds that*

$$s(f_1, f_2) \;=\; p_{\frac{f_2}{f_1}\frac{V_1}{V_2}}(1) \;=\; \tfrac{f_1}{f_2}\, p_{\frac{F_1}{F_2}|\frac{U_1}{U_2}}\left(\tfrac{f_1}{f_2}\,|\,1\right) \;=\; \int_0^\infty \tfrac{f_1 f_2}{u^3}\, p_{V_1}\left(\tfrac{f_1}{u}\right)\, p_{V_2}\left(\tfrac{f_2}{u}\right) du. \quad (16)$$

Moreover, $s(\cdot,\cdot)$ is symmetric and has the following properties:

i) $s(f, f) = const$ *for all* $f > 0$,
ii) $0 \leq s(f_1, f_2) \leq s(f, f) = p_{\frac{V_1}{V_2}}(1)$ *for all* $f_1, f_2, f > 0$.

Note that (16) differs from (13) only by the factor $\frac{f_1 f_2}{u}$ within the integral. Regarding (12) and (14), our similarity measure is not exactly $p_{\frac{F_1}{F_2}|\frac{U_1}{U_2}}(\frac{f_1}{f_2}\,|\,1)$, but a scaled version of it. For multiplicative Gamma noise we have

$$s(f_1, f_2) = \frac{\Gamma(2L)}{\Gamma(L)^2}\,\frac{(f_1 f_2)^L}{(f_1 + f_2)^{2L}} \;=\; \frac{\Gamma(2L)}{\Gamma(L)^2}\left(2 + \frac{f_1}{f_2} + \frac{f_2}{f_1}\right)^{-L}, \qquad f_1, f_2 > 0,$$

with a maximum of $c = p_{\frac{V_1}{V_2}}(1) = \frac{\Gamma(2L)}{\Gamma(L)^2}\,\frac{1}{4^L}$. Fig. 3 shows that for multiplicative Gamma noise we obtain by $s(\cdot,\cdot)/c$ similar histograms as initially for additive Gaussian noise in Fig. 1. Hence, a similar good performance can be expected if applied for nonlocal filtering.

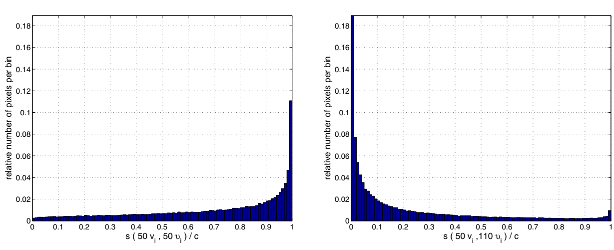

Fig. 3. *Left*: Histogram of $(s(f_i, \widetilde{f}_i)/c)_{i=1}^N$, where f, $\widetilde{f}$ are both constant images of gray value 50 corrupted by multiplicative Gamma noise with $L = 16$. *Right*: Same as on the left, but now $\widetilde{f}$ represents a constant image of gray value 110 corrupted by noise.

5 Weight Definition of Our Nonlocal Filters

For random variables U_i, V_i, F_i, $i = 1, \ldots, N$, fulfilling the multiplicative noise model in Subsection 3.2 with unknown distribution p_{U_i}, the weights can now be defined similarly to (7) by

$$w(i,j) \;=\; \prod_{k \in I} \left(\frac{1}{c}\, s(f_{i+k}, f_{j+k}) \right)^{\frac{g_k}{h}} \;=\; \prod_{k \in I} \left(p_{\frac{f_{j+k}}{f_{i+k}} \frac{V_{i+k}}{V_{j+k}}}(1) \,/\, p_{\frac{V_{i+k}}{V_{j+k}}}(1) \right)^{\frac{g_k}{h}}. \quad (17)$$

As before, $h > 0$ and $g = (g_k)_{k \in I}$ represents a sampled two dimensional Gaussian kernel with mean zero and standard deviation a, which we normalize such that $\sum_{k \in I} g_k = 1$. Besides, the index set I is set to be a squared grid of size $l \times l$ centered at 0 using reflecting boundary conditions for f.

Fig. 4 (top) shows the histograms of the weights (17) for different constant patches corrupted by multiplicative Gamma noise. As visible here, multiplying the values of the similarity measure over a whole patch significantly changes the histograms compared to Fig. 3. Now, the weights of the left histogram are all larger than on the right. Unfortunately, the histogram on the left is no longer maximal at 1. Even worse, weights close to 1 have never been assigned.

To overcome this drawback we propose an additional adaptation of the weights inspired by the implementation of the NL means filter described at [2]. Here, we use that for random variables X, Y and a continuous function b, where $\mathbb{E}(b(Y))$ exists, the *conditional expectation* of $b(Y)$ given $X = x$ is

$$\mathbb{E}(b(Y)|X = x) \;:=\; \int_{-\infty}^{\infty} b(y)\, p_{Y|X}(y|x)\, dy \qquad \forall\, x \text{ with } p_X(x) > 0,$$

see, e.g., [13, p. 168]. In detail, for two sets of random variables $F_{i+k} = U_{i+k}V_{i+k}$, $F_{j+k} = U_{j+k}V_{j+k}$, $k \in I$, we set

$$b_k\left(\tfrac{f_{i+k}}{f_{j+k}}\right) := \left(\tfrac{1}{c}\, p_{\frac{f_{j+k}}{f_{i+k}} \frac{V_{i+k}}{V_{j+k}}}(1)\right)^{\frac{g_k}{h}} = \left(\tfrac{1}{c}\, s(f_{i+k}, f_{j+k})\right)^{\frac{g_k}{h}}$$

and compute for disjoint index sets $i + I$, $j + I$ the conditional expectation

$$\mu := \mathbb{E}\left(\prod_{k \in I} b_k\left(\tfrac{F_{i+k}}{F_{j+k}}\right) \;\middle|\; \left(\tfrac{U_{i+k}}{U_{j+k}} = 1\right)_{k \in I} \right) = \prod_{k \in I} \mathbb{E}\left(b_k\left(\tfrac{F_{i+k}}{F_{j+k}}\right) \;\middle|\; \tfrac{U_{i+k}}{U_{j+k}} = 1 \right).$$

Since $w(i,j)$ is a realization of $\prod_{k \in I} b_k\left(\frac{F_{i+k}}{F_{j+k}}\right)$, the variable μ denotes the value we can expect for $w(i,j)$ if the (non-overlapping) image patches f_{i+I}, f_{j+I} have been generated from the same noise free patch. We can show that

$$\mu \;=\; \prod_{k \in I} \int_0^{\infty} b_k(t)\, p_{\frac{V_{i+k}}{V_{j+k}}}(t)\, dt.$$

using properties of the conditional expectation. For *multiplicative Gamma noise* we obtain by technical computations that

$$\mu = \prod_{k \in I} 4^{Lg_k/h} \frac{\Gamma(2L)}{\Gamma(L)^2} \frac{\Gamma(L(1 + \frac{g_k}{h}))^2}{\Gamma(2L(1 + \frac{g_k}{h}))}.$$

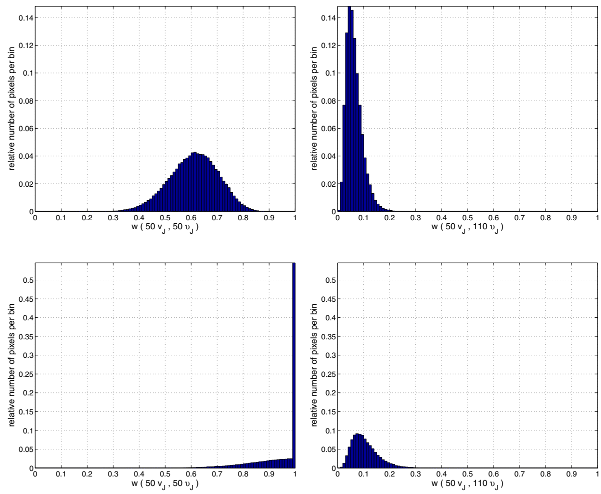

Fig. 4. Histograms of the weights (17) (top) and (18) (bottom) used to compare N different image patches f_I, $\widetilde{f}_I$ ($l = 5$, $a = 1.5$, $h = 1$, $q = 0$). *Left*: Both f_I, $\widetilde{f}_I$ are image patches of gray value 50 corrupted by multiplicative Gamma noise with $L = 16$. *Right*: Same as on the left, but now $\widetilde{f}_I$ is of gray value 110 and corrupted by noise.

Now, we set for $q \in [0, 1)$

$$w_{\mu,q}(i,j) := \begin{cases} 1 & \text{if } w(i,j) \geq \mu, \\ \frac{w(i,j)}{\mu} & \text{if } q\mu \leq w(i,j) < \mu, \\ 0 & \text{otherwise} \end{cases} \qquad \forall\, i,j \in \{1,\dots,N\} \qquad (18)$$

and use these weights in our nonlocal filters deduced from (2). Here, μ is used as an approximation of the true expectation value for all overlapping image patches.

The effect of this additional adaptation compared to (17) is visualized in Fig. 4 (bottom). The histogram for the image patches generated from the same noise free patch has now a significant peak at 1. By setting, e.g., $q = 0.5$ we can additionally achieve that all weights of the right histogram obtain an optimal weight of 0 without effecting the weights of the left histogram.

As usually done, we finally restrict the number of patches being compared to a so-called *similarity window*. Thus, we set all weights $w(i,j)$, $w_{\mu,q}(i,j)$ automatically to zero if pixel j is outside of a squared image region of size $\omega \times \omega$ centered at pixel i. This reduces the computational costs as well as the risk of falsely assigning nonzero weights to a large number of patches.

Updating the Similarity Neighborhoods

In [6] Deledalle et al. suggest to refine the weights of their nonlocal filters iteratively using the former result $u^{(r-1)}$. To obtain $u^{(r)}$, the filter is again applied to the initial noisy image using the new weights. The idea for this updating scheme was taken from [12]. In the following, we apply a variant of this updating strategy, where we perform only one updating step. For this second step we use within the similarity windows for $i, j = 1, \ldots, N$, $i \neq j$ the weights

$$\widetilde{w}_{i,j}(u^{(1)}) = \exp\left(-\frac{1}{d} \sum_{k \in \widetilde{I}} \widetilde{g}_k \; K_{\text{sym}}\left(p_{F_{i+k}|U_{i+k}}(\,\cdot\,|u^{(1)}_{i+k}), p_{F_{j+k}|U_{j+k}}(\,\cdot\,|u^{(1)}_{j+k}) \right) \right)$$

and set $\widetilde{w}_{i,i}(u^{(1)}) = \max_j \widetilde{w}_{i,j}(u^{(1)})$. Here, $d > 0$ and $\widetilde{g} = (\widetilde{g}_k)_{k \in \widetilde{I}}$ is again a sampled two dimensional Gaussian kernel with mean zero, but with standard deviation $\widetilde{a}$. As before, $\widetilde{g}$ is normalized such that $\sum_{k \in \widetilde{I}} \widetilde{g}_k = 1$. Moreover, $\widetilde{I} = \widetilde{l} \times \widetilde{l}$ may vary from I. Usually, we choose $\widetilde{a} < a$ and $\widetilde{l} < l$. Furthermore,

$$K_{\text{sym}}(p_X, p_Y) := \int_{-\infty}^{\infty} (p_X(t) - p_Y(t)) \ln\left(\frac{p_X(t)}{p_Y(t)}\right) dt$$

denotes the *symmetric Kullback-Leibler divergence* of p_X, p_Y. If we assume that $p_{U_i}(x) > 0$ for all $x \geq 0$, we can show using (5) that the sought symmetric Kullback-Leibler divergence for multiplicative Gamma noise is given by

$$K_{\text{sym}}\left(p_{F_i|U_i}(\,\cdot\,|u^{(1)}_i), p_{F_j|U_j}(\,\cdot\,|u^{(1)}_j) \right) = L \frac{(u^{(1)}_i - u^{(1)}_j)^2}{u^{(1)}_i u^{(1)}_j} \qquad \text{for } u^{(1)}_i, u^{(1)}_j > 0.$$

6 Numerical Results

Finally, we present two examples demonstrating the good performance of our novel nonlocal filters for images corrupted by multiplicative Gamma noise. The implementation was done with MATLAB and the parameters were chosen to obtain the best visual results. Note that all images, especially the noisy one, are displayed in the gray scale of the original image to have a consistent coloring for each example. To this purpose, all image values outside of the range of the original image are projected on this range.

For our first example we use the same test image as in [14, Fig. 6]. Obviously, our reconstructions in Fig. 5 (bottom middle and right) are superior to the result by the I-divergence - TV method at top left. Moreover, the difference of applying (6) with weights $w(i, j)$ or $w_{\mu,q}(i, j)$ is illustrated. By using $w_{\mu,q}(i, j)$ instead of $w(i, j)$ more noise has been removed, especially in the background. Moreover, an appropriate value q helps to improve the contrast, e.g., visible at the camera, and leads to sharper edges and contours. By the final updating step used for Fig. 5 (bottom right) we further improved the contrast and small amounts of possibly remained noise are finally removed.

Our second example in Fig. 6 shows our result for the noisy image in [8, Fig. 8]. For a better comparison we included its peak signal to noise ratio (PSNR) and mean absolute-deviation error (MAE) as, e.g, defined in [8]. Obviously, our result is superior or at least competitive to the results obtained by various methods in [8, Fig. 8]. There, the best result was obtained by the proposed hybrid multiplicative noise removal method, which combines variational and sparsity-based shrinkage methods involving curvelets and TV regularization.

Fig. 5. *Top*: Original image with values in $[0, 255]$ (left), corrupted version by multiplicative Gamma noise with $L = 4$ (middle) and restored image by the I-divergence - TV model as presented in [14]. *Bottom*: Results by our new nonlocal filter (6) using just (17) with $l = 7$, $\omega = 29$, $a = 1.5$, $h = 1$ (left), using (18) with $q = 0.35$ (middle) and after an additional updating step with $\widetilde{l} = 3$, $\widetilde{a} = 0.5$, $d = 0.25$ (right).

Fig. 6. *Left*: Original image of the French city of Nîmes (512×512) with values in $[1, 256]$, which has been corrupted by multiplicative Gamma noise with $L = 4$ in [8, Fig. 8]. *Right*: Restoration result by our nonlocal filter (6) applied to the noisy image using (18) and an additional updating step with $l = 7$, $\omega = 29$, $a = 2$, $h = 0.5$, $q = 0.7$, $\widetilde{l} = 5$, $\widetilde{a} = 1$, $d = 0.1$ (PSNR = 26.01, MAE = 8.60).

Acknowledgment. The authors would like to thank the authors of [8], in particular Mila Nikolova, for kindly providing the initial data used in Fig. 6.

References

1. Aubert, G., Aujol, J.-F.: A variational approach to removing multiplicative noise. SIAM Journal on Applied Mathematics 68(4), 925–946 (2008)
2. Buades, A., Coll, B., Morel, J.-M.: Online demo: Non-local means denoising, `http://www.ipol.im/pub/algo/bcm_non_local_means_denoising`
3. Buades, A., Coll, B., Morel, J.-M.: A non-local algorithm for image denoising. In: IEEE Conf. on CVPR, vol. 2, pp. 60–65 (2005)
4. Buades, A., Coll, B., Morel, J.-M.: Image denoising methods. A new nonlocal principle. SIAM Review 52(1), 113–147 (2010)
5. Coupé, P., Hellier, P., Kervrann, C., Barillot, C.: Nonlocal means-based speckle filtering for ultrasound images. IEEE Trans. Image Process. 18(10), 2221–2229 (2009)
6. Deledalle, C.-A., Denis, L., Tupin, F.: Iterative weighted maximum likelihood denoising with probabilistic patch-based weights. IEEE Trans. Image Process. 18(12), 2661–2672 (2009)
7. Deledalle, C.-A., Tupin, F., Denis, L.: Poisson NL means: Unsupervised non local means for Poisson noise. In: Proceedings of IEEE International Conference on Image Processing, pp. 801–804 (2010)
8. Durand, S., Fadili, J., Nikolova, M.: Multiplicative noise removal using L1 fidelity on frame coefficients. J. Math. Imaging Vision 36(3), 201–226 (2010)
9. Grimmett, G.R., Stirzaker, D.R.: Probability and random processes, 3rd edn. Oxford University Press, Oxford (2001)
10. Kervrann, C., Boulanger, J., Coupé, P.: Bayesian non-local means filter, image redundancy and adaptive dictionaries for noise removal. In: Sgallari, F., Murli, A., Paragios, N. (eds.) SSVM 2007. LNCS, vol. 4485, pp. 520–532. Springer, Heidelberg (2007)
11. Matsushita, Y., Lin, S.: A probabilistic intensity similarity measure based on noise distributions. In: IEEE Conf. Computer Vision and Pattern Recognit. (2007)
12. Polzehl, J., Spokoiny, V.: Propagation-separation approach for local likelihood estimation. Probability Theory and Related Fields 135(3), 335–362 (2006)
13. Rohatgi, V.K.: An Introduction to Probability Theory and Mathematical Statistics. John Wiley & Sons, Inc., Chichester (1976)
14. Steidl, G., Teuber, T.: Removing multiplicative noise by Douglas-Rachford splitting methods. J. Math. Imaging Vision 36(2), 168–184 (2010)
15. Teuber, T., Lang, A.: A new similarity measure for nonlocal filtering in the presence of multiplicative noise. University of Kaiserslautern (preprint, 2011)
16. Wiest-Daesslé, N., Prima, S., Coupé, P., Morrissey, S.P., Barillot, C.: Rician noise removal by non-local means filtering for low signal-to-noise ratio MRI: Applications to DT-MRI. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part II. LNCS, vol. 5242, pp. 171–179. Springer, Heidelberg (2008)